

What Is Claimed:

1. A method for generating geographic coordinate information comprising:
receiving at least one term that specifies an address;
accessing a table defining coordinate information for ranges of addresses to find
an intersection of sets of rows in the table that correspond to the at least one term; and
reading geographic coordinate information from the table at the intersection of the
sets of rows in the table.
2. The method of claim 1, wherein the at least one term includes a plurality
of terms.
3. The method of claim 1, wherein reading the geographic coordinate
information includes:
reading starting geographic coordinate values from the table;
reading ending geographic coordinate values from the table;
reading a starting street number from the table;
reading an ending street number from the table; and
calculating the geographic coordinate information based on an interpolation of an
actual street number relative to the starting and ending street numbers and the starting and
ending geographic coordinate values.
4. The method of claim 1, wherein columns of the table include fields that
correspond to the at least one term that specifies the address.

5. The method of claim 4, wherein the fields in the table include at least a state field, a city field, a zip code field, and a street field.
6. The method of claim 1, wherein the coordinate information is defined as latitude and longitude values.
7. The method of claim 4, wherein the intersection of the sets of rows in the table that correspond to the at least one term is determined by successively locating first sets of rows corresponding to one of the terms and then locating a next set of rows as a subset of the first set of rows corresponding to a next one of the terms.
8. The method of claim 7, wherein the table is successively sorted by a plurality of table columns, the sorting being performed in a sorting order based on an order of the fields used to locate the blocks of rows.
9. The method of claim 8, wherein sections of the sorted columns of the table are stored as a value and an associated beginning and ending row of the value to thereby reduce a stored size of the table.
10. The method of claim 1, wherein the receiving terms that specify an address includes:
extracting the address from a text document.

11 The method of claim 10, wherein extracting the address from the text document further includes:

- identifying possible address terms in the text document based on predetermined rules,
- verifying that the identified possible address terms are address terms by comparing the address terms to fields in the table, and
- examining a relative position of the verified possible address terms in the document to determine whether the verified possible address terms form a valid address.

12. A device comprising:

- means for receiving at least one term that specifies an address;
- means for accessing a table defining coordinate information for ranges of addresses to find an intersection of sets of rows in the table that correspond to the at least one term; and
- means for reading geographic coordinate information from the table at the intersection of the sets of rows in the table.

13. The device of claim 12, wherein the means for receiving further comprises:

- means for identifying possible address terms from a document based on predetermined rules;
- means for verifying that the identified possible address terms are address terms by comparing the address terms to the table; and

means for examining a relative position of the verified possible address terms in the document to determine whether the verified possible address terms form a valid address.

14. A system for geocoding postal addresses comprising:
a table including a plurality of rows that each correspond to a range of one or more addresses, each of the rows including a plurality of fields that define the row; and
a geocoding component configured to generate geographic coordinate information for a received address specified by one or more terms that correspond to the fields by locating at least one row in the table that corresponds to an intersection of a number of sets of rows defined by the terms in the received address.

15. The system of claim 14, wherein the geographic coordinate information is generated by interpolating a street address within a range of street addresses specified in the at least one row.

16. The system of claim 14, wherein the plurality of fields include a state field, a zip code field, and a street field.

17. The system of claim 16, wherein the plurality of fields further include a county field, a city field, a street base name field, and a parity field.

18. The system of claim 17, wherein the plurality of fields further include a starting street number field, an ending street number field, starting geographic coordinate information fields, and ending geographic coordinate information fields.

19. The system of claim 14, wherein columns of the table correspond to the plurality of fields.

20. The system of claim 14, wherein the geographic coordinate information is defined as latitude and longitude values.

21. A method for extracting addresses from a document, the method comprising:

identifying possible address terms based on predetermined rules;
verifying that the identified possible address terms are address terms by
comparing the address terms to a table containing known addresses; and
examining a relative position of the verified possible address terms in the
document to determine whether the verified possible address terms form a valid address.

22. The method of claim 21, wherein the predetermined rules include a comparison of the possible address terms to a known list of possible terms.

23. The method of claim 21, wherein the predetermined rules include a comparison of capitalization in the possible address terms to capitalization consistent with actual address terms.

24. The method of claim 21, wherein the predetermined rules include a verification that the possible address terms that define street names are preceded by a number.

25. The method of claim 21, further comprising normalizing the possible address terms to a standardized set of address terms.

26. The method of claim 21, further comprising:
omitting prefixes and suffixes when verifying that the identified possible address terms are address terms.

27. A computing device comprising:
a memory configured to store a table containing valid addresses; and
a processor operatively coupled to the memory, the processor executing programming instructions that cause the processor to extract addresses from a document by:
identifying possible address terms in the document based on predetermined rules,

verifying that the identified possible address terms are address terms by comparing the address terms to the table, and

examining a relative position of the verified possible address terms in the document to determine whether the verified possible address terms form a valid address.

28. A computer-readable medium useable by a processor containing a data structure comprising:

a table including a plurality of rows that each correspond to a range of one or more addresses, each of the rows including a plurality of fields that define the row, the fields including at least a street address field and a zip code field, the table being organized such that a row in the table corresponding to a particular address is accessed as an intersection of a plurality of sets of rows defined by values of a received postal address.